

AD-A111 219

MARYLAND UNIV COLLEGE PARK

F/G 12/2

ERRORS IN INSPECTION AND GRADING: DISTRIBUTIONAL ASPECTS OF SCR--ETC(U)

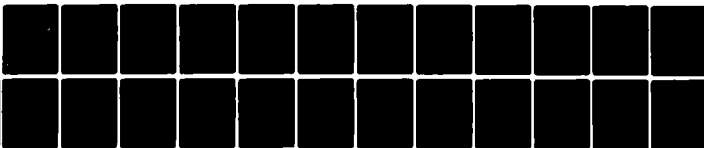
FEB 82 S KOTZ, N L JOHNSON

N00014-81-K-0301

NL

UNCLASSIFIED

1 OF 1
AD A111 219



END

DATE

FORMED

3-82

DTIC

3

AD A111219

ERRORS IN INSPECTION AND GRADING:
DISTRIBUTIONAL ASPECTS OF
SCREENING AND HIERARCHICAL SCREENING

by

Samuel Kotz
University of Maryland

and

Norman L. Johnson
University of North Carolina at Chapel Hill

STIC
FEB 22 1982
E

ABSTRACT

Continuing the studies of Johnson *et al* (1980) and Johnson and Kotz (1981), further distributions arising from models of errors in inspection and grading of samples from finite, possibly stratified lots are obtained. Screening, and hierarchical screening forms of inspection are also considered, and the effects of errors on the advantages of these techniques assessed.

DTIC FILE COPY

This document has been approved
for public release and sale; its
distribution is unlimited.

DTIC
COPY
INSPECTED
2

| | |
|--------------------|-------------------------------------|
| Accession For | |
| NTIS GRA&I | <input checked="" type="checkbox"/> |
| DTIC TAB | <input type="checkbox"/> |
| Unannounced | <input type="checkbox"/> |
| Justification | |
| By _____ | |
| Distribution/ | |
| Availability Codes | |
| for | |
| Dist _____ | |
| A | 23 |

219

1. INTRODUCTION

For convenience, we first summarize some results obtained by Johnson et al. (1980) and Johnson & Kotz (1981).

A random sample of size n is chosen (without replacement) from a lot of size N , which contains D defective (or nonconforming) members. Suppose that on inspection of items in the sample, the probability that a defective item will be classified as defective is p , while the probability that a non-defective item will (erroneously) be classified as "defective" is p' .

If Y denotes the actual unknown number of defective items in the sample, the distribution of Y is hypergeometric with parameters (n, D, N) . If X denote the number (among these Y) which are (correctly) classified as defective, and X' the number (among the $(n-Y)$ nondefective items in the sample) which are (incorrectly) classified as "defective" then, conditionally on Y , the variables X and X' are independent binomial with parameters (Y, p) and $(n-Y, p')$ respectively. Averaging over the distribution of Y , the distribution of

$$Z = X + X',$$

the total number of items described as "defective" after inspection of the sample, can be formally expressed as

$$\text{Binomial}(Y, p) * \text{Binomial}(n-Y, p') \hat{=} \text{Hypergeometric}(n, D, N) \quad (1)$$

The symbol $*$ stands for convolution, and the symbol $\hat{=}$ indicates the "corresponding" operation with Y distributed as "(see, e.g. Johnson & Kotz (1969, Chapter 8)). Distribution (1) is a mixture of convolutions of two binomial distributions. The r -th descending factorial moment of Z is

$$\mu_{(r)}(Z) = E(Z^{(r)}) = \frac{n^{(r)}}{N^{(r)}} \sum_{j=0}^r \binom{r}{j} p^j p'^{r-j} D^{(j)} (N-D)^{(r-j)} \quad (2)$$

DISCLAIMER NOTICE

**THIS DOCUMENT IS BEST QUALITY
PRACTICABLE. THE COPY FURNISHED
TO DTIC CONTAINED A SIGNIFICANT
NUMBER OF PAGES WHICH DO NOT
REPRODUCE LEGIBLY.**

where $Z^{(r)} = Z(Z-1) \dots (Z-r+1)$. In particular

$$E(Z) = n\bar{p}/N \quad (3.1)$$

$$\text{var}(Z) = n\bar{p}(1-\bar{p}) - \frac{n(n-1)}{N-1} \cdot \frac{D}{N} \left(1 - \frac{D}{N}\right) (p-p')^2 \quad (3.2)$$

where $\bar{p} = \{Dp + (N-D)p'\}N^{-1}$ is the probability that a single individual, chosen at random from the lot, will be described as "defective" after inspection (whether it is really defective or not). Tables of the distribution of Z for $n=10$ with $p = 0.75(0.05)0.95$; $p' = 0(0.025)0.1$; $N = 100$ and $D = 5, 10, 20$; $N = 200$ and $D = 10, 20, 40$ are presented in Johnson and Kotz (1981). If $p' = 0$, we have the situation described in Johnson et al. (1980) and (1) becomes a hypergeometric-binomial distribution. For fixed values of $D/N = \lambda$ say, the distribution (1) is quite sensitive to changes in p and p' , but not to changes in D and N unless n/N (the sampling fraction) is large. It is easy to see that as $D, N \rightarrow \infty$ (with $D/N = \lambda$) the distribution of Z tends to binomial with parameters $(n, \lambda p + (1-\lambda)p')$.

Johnson and Kotz (1981) also, *inter alia*, consider lots divided into k strata Π_1, \dots, Π_k of sizes N_1, \dots, N_k (with $\sum_{j=1}^k N_j = N$) and suppose that the probability of an individual from the j -th stratum being classified as defective is p_j . (The situation considered at the beginning of this section corresponds to $k=2$, $N_1=D$, $N_2=N-D$, $p_1 = p$, $p_2 = p'$.) The distribution of the total number (Z) classified as defective is (in an obvious notation)

$$\sum_{j=1}^k \text{Binomial}(Y_j, p_j) \sim \text{multivariate Hypergeometric}(n, \underline{N}, N)$$

($\underline{Y} = (Y_1, \dots, Y_k)$ is the vector of numbers from Π_1, \dots, Π_k in the sample - $Y_1 + \dots + Y_k = n$) (4)

The r -th factorial moment of Z is

$$\frac{n^{(r)}}{N^{(r)}} = \sum_r \left\{ \begin{matrix} r \\ r_1, r_2, \dots, r_k \end{matrix} \right\} \prod_{j=1}^k \left\{ \begin{matrix} r_j \\ p_j N_j \end{matrix} \right\} \quad (5)$$

where \sum_r' denotes summation subject to $r_1 + \dots + r_k = r$ and (r_1, r_2, \dots, r_k) is the multinomial coefficient $r! / \left(\prod_{j=1}^k r_j! \right)$.

In particular, with $\bar{p} = N^{-1} \sum_{j=1}^k N_j p_j$ (again, denoting the probability that an individual chosen at random is classified as "defective")

$$E(Z) = n\bar{p} \quad (6.1)$$

$$\text{var}(Z) = n\bar{p}(1-\bar{p}) - \frac{n(n-1)}{n(N-1)} \sum_{j=1}^k N_j (p_j - \bar{p})^2 \quad (6.2)$$

(Note that $\text{var}(Z)$ is usually less than the "binomial" value $n\bar{p}(1-\bar{p})$.)

2. GRADING

Formulae (4)-(6) can be regarded as generalizations of (1)-(3). Further useful generalization is obtained by considering judgement not to be restricted to "defective" or "non-defective" but to include assignment to one of several categories - as would be the case, if product were being graded in terms of quality and/or size with regard to marketing. (One aspect of this situation is the multivariate topic of disjunctive analysis, but here we are concerned with the consequences rather than the methods of assignment.)

We will analyze a situation in which the aim of judgement is assign an individual to one of s classes C_1, \dots, C_s . We denote the probability that an individual, who really belongs to C_j will be assigned to C_i by p_{ij} , with, of course $\sum_{i=1}^s p_{ij} = 1$. Still further generalization is possible by introducing stratification within each class. This leads to straightforward, but notationally complicated elaboration, and will not be pursued here.

Let Y_j denote the number of individuals belonging to C_j , in a random sample of size n for a lot of size N containing N_j individuals in C_j ($j=1, \dots, s$; $\sum_{j=1}^s N_j = N$); and let Z_{ij} denote the number, among these Y_j , assigned to

C_i ($i=1, \dots, s$). The \underline{Y} has a multivariate hypergeometric distribution with parameters $(n; \underline{N}; N)$, so that

$$P(\underline{Y}) = \left\{ \prod_{j=1}^s \binom{N_j}{Y_j} \right\} / \binom{N}{n} \quad \left(\sum_{j=1}^s Y_j = n \right) \quad (7)$$

Also, given \underline{Y} , $\underline{Z}_j = (Z_{1j}, \dots, Z_{sj})$ has a multinomial distribution with parameters (Y_j, \underline{P}_j) where $\underline{P}_j = (P_{1j}, \dots, P_{sj})$, and

$$P(\underline{Z}_j | \underline{Y}) = Y_j! \prod_{i=1}^s (P_{ij}^{Z_{ij}} / Z_{ij}!) \quad \left(\sum_{i=1}^s Z_{ij} = Y_j \right) \quad (8)$$

and $\underline{Z}_1, \dots, \underline{Z}_s$ are mutually independent. It follows that the joint distribution of all the \underline{Z} 's, $\underline{Z} = (\underline{Z}_1, \dots, \underline{Z}_s)$ is

$$P(\underline{Z}) = P(\underline{Z} | \underline{Y}) P(\underline{Y}) = \binom{N}{n}^{-1} \prod_{j=1}^s \left\{ \binom{N_j}{Y_j} \prod_{i=1}^s (P_{ij}^{Z_{ij}} / Z_{ij}!) \right\} \quad (9)$$

where $Z_{ij} = \sum_{i=1}^s Z_{ij}$. (Note that the Y_j 's are determined by the Z_{ij} 's.)

Formally we can write the distribution of \underline{Z} as

$$\prod_{j=1}^s \text{Multinomial}(Y_j, \underline{P}_j) \wedge \text{Multivariate Hypergeometric}(n, \underline{N}, N) \quad (10)$$

and it might be called a "multivariate hypergeometric-multinomial distribution."

The joint factorial moments of \underline{Z} can be obtained from

$$\begin{aligned} E \left[\prod_{j=1}^s \prod_{i=1}^s Z_{ij}^{r_{ij}} \right] &= E_{\underline{Y}} \left[E \left[\prod_{j=1}^s \prod_{i=1}^s Z_{ij}^{r_{ij}} | \underline{Y} \right] \right] = E \left[\prod_{j=1}^s Y_j^{(r_{..j})} \prod_{i=1}^s P_{ij}^{r_{ij}} \right] \\ &= \frac{n(r_{..})}{N(r_{..})} \prod_{j=1}^s \left\{ \binom{N_j}{N_j} \prod_{i=1}^s P_{ij}^{r_{ij}} \right\} \quad (11) \end{aligned}$$

where $r_{.j} = \sum_{i=1}^s r_{ij}$; $r_{..} = \sum_{i=1}^s \sum_{j=1}^s r_{ij}$. In particular, with $i \neq i', j \neq j'$

$$E(Z_{ij}) = nN^{-1}N_j P_{ij} \quad (12.1)$$

$$E(Z_{ij}Z_{i',j'}) = n^{(2)}N_j^{(2)}P_{ij}P_{i',j'}/N^{(2)} \quad (12.2)$$

$$E(Z_{ij}Z_{i',j'}) = n^{(2)}N_jN_{j'}P_{ij}P_{i',j'}/N^{(2)} \quad (12.3)$$

whence

$$\text{cov}(Z_{ij}, Z_{i',j'}) = -nN_j\{(N-n)N_{j'} + N(n-1)\}P_{ij}P_{i',j'}N^{-2}(N-1)^{-1} \quad (13.1)$$

$$\text{cov}(Z_{ij}, Z_{i',j'}) = -nN_jN_{j'}(N-n)P_{ij}P_{i',j'}N^{-2}(N-1)^{-1} \quad (13.2)$$

Using these results we can find the covariance between $Z_{i.} = \sum_{j=1}^s Z_{ij}$ and $Z_{i'.} = \sum_{j=1}^s Z_{i'j}$ - that is the total number of individuals assigned to $C_i, C_{j'}$, respectively.

$$\text{cov}(Z_{i.}, Z_{i'.}) = -n\{(N-n)\bar{P}_i\bar{P}_{i'} + (n-1)(\bar{P}_i\bar{P}_{i'})\} \quad (14)$$

where

$$\bar{P}_i = N^{-1} \sum_{j=1}^s N_j P_{ij}; \quad (\bar{P}_i\bar{P}_{i'}) = N^{-1} \sum_{j=1}^s N_j P_{ij} P_{i'j}$$

(\bar{P}_i is the probability that an individual chosen at random is assigned to C_i , ($\bar{P}_i\bar{P}_{i'}$) is the probability that an individual chosen at random would be assigned to C_i on one judgement and $C_{i'}$ on another.). The marginal distribution of $Z_{i.}$ is like (4) with k replaced by s and p_i by $P_{ij.}$, so we can use the corresponding formulae for moments.

3. GROUP SCREENING

Further interesting distributions arise in connection with "group screening" (Dorfman (1943)), in which groups of units can be tested for the existence of one or more defective units among them. This can be practicable, for example, when testing liquids for presence of contaminants, and is then suggested as a possible way of reducing the average total amount of testing.

Suppose that material from n_1 units is mixed and tested for presence of "defective" material. If a negative result ("no defectives") is obtained, no further action is taken, but if there is a positive result, each unit is tested separately.

Let p_1, p'_1 denote the probabilities of obtaining correct or incorrect positive results, respectively, at the first test. As before, p, p' denote the probabilities of correct or incorrect positive results, respectively, when units are tested individually; D, N denote the number of defective units and the total number of units in the population respectively, and Y denotes the actual number of defective units among the n_1 tested.

The overall probability of obtaining a positive result on the first test is

$$\{1 - P_0(n_1)\}p_1 + P_0(n_1)p'_1 = p_1 - (p_1 - p'_1)P_0(n_1) \quad (15)$$

where $P_0(n_1) = (N-D) \binom{n_1}{0} / \binom{N}{n_1}$ is the probability that the sample contains no defective items.

The distribution of Z can be represented as

$$(\text{Binomial}(WY, p) * \text{Binomial}(W(n_1 - Y), p')) \overset{\Delta}{=} \text{Hypergeometric}(n_1, D, N) \quad (16)$$

where W is an indicator variable, defined by

$$W = \begin{cases} 1 & \text{if the first test gives a positive result} \\ 0 & \text{otherwise} \end{cases}$$

Denote that

$$P(W=1 | Y>0) = p_1 ; \quad P(W=1 | Y=0) = p'_1 \quad (17)$$

An explicit formula for $P(Z=z)$ is

$$P(Z=z) = E_{W,Y} \left[\sum_x \binom{WY}{x} p^x (1-p)^{WY-x} \binom{W(n_1-Y)}{z-x} p'^{z-x} (1-p')^{W(n_1-Y)-z+x} \right]$$

where \sum_x denotes summation over $(0, z - W(n_1 - Y)) \leq x \leq \min(WY, z)$.

We have

$$E(Z^{(r)} | W, Y) = \sum_{i=0}^r \binom{r}{i} (WY)^{(i)} \{W(n_1 - Y)\}^{(r-i)} p^i p'^{r-i}$$

Noting (17) and

$$E(Y^{(i)} (n_1 - Y)^{(r-i)} | Y > 0) P(Y = 0) = \begin{cases} n_1^{(r)} D^{(i)} (N-D)^{(r-i)} / N^{(r)} & \text{for } i > 0 \\ n_1^{(r)} [(N-D)^{(r)} \{N^{(r)}\}^{-1} - p_0(n_1)] & \text{for } i = 0 \end{cases} \quad (18.1)$$

$$E(Y^{(i)} (n_1 - Y)^{(r-i)} | Y = 0) P(Y = 0) = \begin{cases} 0 & \text{for } i > 0 \\ n_1^{(r)} p_0(n_1) & \text{for } i = 0 \end{cases} \quad (18.2)$$

we find

$$E(Z^{(r)}) = n_1^{(r)} \{p_1 \{N^{(r)}\}^{-1} \sum_{i=0}^r \binom{r}{i} D^{(i)} (N-D)^{(r-i)} p^i p'^{r-i} - (p_1 - p'_1) p'^r p_0(n_1)\} \quad (19.1)$$

In particular

$$\begin{aligned} E(Z) &= n_1 \{p_1 \bar{p} - (p_1 - p'_1) p' p_0(n_1)\} \\ &= n_1 (p_1 \bar{p} - p_{(1)} p') \end{aligned} \quad (19.2)$$

where

$$p_{(1)} = (p_1 - p'_1) p'(n_1)$$

and

$$\begin{aligned} \text{var}(Z) &= n_1^{(2)} \left[p_1 \left\{ \bar{p}^2 - \frac{D(p^2 - \bar{p}^2) + (N-D)(p'^2 - \bar{p}^2)}{N(N-1)} \right\} - p_{(1)}^2 p'^2 \right] \\ &\quad + n_1 (p_1 \bar{p} - p_{(1)} p') - n_1^2 (p_1 \bar{p} - p_{(1)} p')^2 \end{aligned} \quad (19.3)$$

If there is no possibility of a "false positive" (so that $p'_1 p' = 0$) we find

$$\begin{aligned} E(Z) &= n_1 p_1 p D / N \\ \text{Var}(Z) &= \frac{n_1 p_1 p D}{N} \left(1 - \frac{p_1 p D}{N} \right) + \frac{n_1 (n_1 - 1) p D}{N^2 (N-1)} \{N(D-1) - (N-1) D p_1\} \end{aligned}$$

In general, it is to be expected (and hoped) that $p_1 > p'_1$ just as $p > p'$, since we would expect (hope) that the probability of correct decision would exceed that of incorrect decision. It may well happen that $p_1 < p$ since detection of a defective may be more difficult with the mixture of material from separate units. More complicated distributions will be obtained if it is supposed that p_1 depends on the value of Y (the number of defective units). It does not seem unreasonable to suppose that p_1 might increase with Y .

The effectiveness of the screening procedure is measured by the three quantities

$$(i) \text{ Probability of correct classification for defectives} = p_1 p \quad (20.1)$$

$$(ii) \text{ Probability of correct classification for nondefective} =$$

$$P_0^*(n_1)(1-p'_1 p') + (1-P_0^*(n_1))(1-p_1 p') = 1 - (p_1 - P_{(1)}^*)p'$$

where

$$P_{(1)}^* = (P_1 - p'_1)P_0^*(n_1); \text{ and } P_0^*(n_1) = (N-D-1)^{(n_1-1)} / (N-1)^{(n_1-1)} \quad (20.2)$$

is the probability that the sample contains no defectives, given that one member of the sample is nondefective.

$$(iii) \text{ Expected number of tests} = 1 + n_1 [P_0(n_1)p'_1 + \{1-P_0(n_1)\}p_1] \quad (20.3)$$

$$= 1 + n_1 (p_1 - P_{(1)})$$

of course, the larger the values of (i) and (ii), and the smaller the value of (iii), the better.

From (20.1) it is clear that the probability of correct classification of a defective is *decreased* by the screening process. (since $p_1 p < p$). The value of screening must therefore come from increased correct classification of nondefectives or reduction in the expected number of tests. Table 1 contains some relevant numerical information. In the absence of screening, n_1 tests would be necessary, so

$$1 - \{1 + n_1\} p_1 - P_{(1)} / n, = 1 - n_1^{-1} - p_1 + P_{(1)}$$

indicates the proportionate saving from screening, and this is given (as a percentage) in the last column of Table 1. We have taken $N = \infty$, so that $\omega = D/N$ is to be interpreted as proportion defective, because the values do not depend greatly on N . As N decreases, $P_{(1)}$ and $P_{(1)}^*$ decrease so that both the proportionate saving in number of tests from screening and the probability of correct identification of nondefectives decrease, (though not substantially, unless N is quite small). It is to be noted that screening *increases* the probability of correct identification of nondefectives.

4. HIERARCHAL GROUP SCREENING

Sometimes additional saving in the expected number of tests, and improved accuracy in classification, can be attained by using two-or more-stage screening - that is, hierarchal screening. For simplicity, we will consider two-stage procedures, with a first-stage sample of size $n_1 = hn_2$. (Generalization to more than two stages follows similar lines). If a positive results is obtained for the combined sample, it is split into h subsamples, each of size n_2 and each is then treated as in Section 3. Letting p_2, p_2' denote the respective probabilities of correct and incorrect positive results when testing each of the second stage (size n_2) subsamples, the three quantities measuring effectiveness are:

- (i) Probability of correct classification for defectives = $p_1 p_2 p$ (21.1)
- (ii) Probability of correct classification for nondefectives

$$= P_0^*(n_1)(1 - p_1' p_2' p) + (1 - P_0^*(n_1))(1 - p_1) + (P_0^*(n_2) - P_0^*(n_1))p_1(1 - p_2 p')$$

$$+ (1 - P_0^*(n_2))p_1(1 - p_2 p')$$

$$\begin{aligned}
 &= 1 - p_1 p_2 p' + P_0^*(n_1)(p_1 - p_1') p_2' p' + P_0^*(n_2) p_1 (p_2 - p_2') p' \\
 &= 1 - p_1 p_2 p' + (P_{(1)}^* p_2' + P_{(2)}^* p_1) p' \quad (21.2)
 \end{aligned}$$

where $P_{(2)}^* = P_0^*(n_2)(p_2 - p_2')$

(Note that $P_0^*(n_2) - P_0^*(n_1)$ is the probability that a random sample of n_1 , known to contain at least one nondefective, also contains at least one defective, but a randomly chosen subsample of size n , containing at least one nondefective, in fact contains no defectives.)

(iii) Expected number of tests

$$\begin{aligned}
 &= 1 + h(p_1 - p_{(1)}) + n_1 [P_0^*(n_1) p_1' p_2' + (1 - P_0^*(n_2)) p_1 p_2 + (P_0^*(n_2) - P_0^*(n_1)) p_1 p_2'] \\
 &= 1 + h(p_1 - p_{(1)}) + n_1 (p_1 p_2 - P_{(1)}^* p_2' - P_{(2)}^* p_1) \quad (21.3)
 \end{aligned}$$

where $P_{(2)}^* = P_0^*(n_2)(p_2 - p_2')$.

The proportional reduction in expected number of tests is

$$1 - p_1 p_2 - n_2^{-1} (p_1 - p_{(1)}) + P_{(1)}^* p_2' + P_{(2)}^* p_1 \quad (21.4)$$

Values given by (21.2) and (21.4) are shown in Table 2.

From (21.1) we see that the probability of correct classification of a defective is decreased more than for simple screening (cf. (20.1)). As some compensation, the probability of correct classification of nondefectives is higher.

As in the case of simple screening the advantages of a screening procedure are greater when the proportion (ω) of defectives is smaller. The effect of finite lot size (N) is to decrease $P_{(1)}$, $P_{(2)}$, $P_{(1)}^*$, $P_{(2)}^*$. From equations (21) it can be seen that this will

- (i) not affect the probability of correct classification of defectives
- (ii) decrease the probability of correct classification of nondefectives

(iii) decrease the expected number of tests.

Some analysis of the distribution of the total number (Z) of items classified as "defectives" by this hierarchical screening procedure is given in the Appendix. Here we just give the expected value

$$E(Z) = n_1 [p_1 p_2 (p - p') DN^{-1} + (p_1 p_2 - p_{(2)} p_1 - p_{(1)} p_2') p'] \quad (22)$$

ACKNOWLEDGEMENT

Samuel Kotz's work was supported by the U.S. Office of Naval Research under Contract N00014-81-K-0301. Norman L. Johnson's work was supported by the National Science Foundation under Grant MCS-8021704.

REFERENCES

- Dorfman, R. (1945). The detection of defective members of large populations. *Ann. Math. Statist.* 14, 430-440.
- Johnson, N.L. and Kotz, S. (1969). *Distributions in Statistics--Discrete Distributions*. New York: John Wiley and Sons.
- Johnson, N.L. and Kotz, S. (1981). Faulty inspection distributions -- some generalizations. (To be published in *Proc. of ONF/ARO Reliability Workshop*, April 1981) *Institute of Statistics Mimeo Series #1735*, University of North Carolina at Chapel Hill.
- Johnson, N.L., Kotz, S. and Sorkin, H.L. (1980). Faulty inspection distributions. *Commun. Statist.* A9(9), 917-922.
- Johnson, N.L. and Leone, F.C. (1977). *Statistics and Experimental Design: In Engineering and the Physical Sciences*, 2nd Edition. New York: John Wiley and Sons. (pp 914-915).

Table 1: Simple Screening

- (i) Probability of correct classification of defective is $p_1 p$
(ii) These tables correspond to $N = \infty$. Values of $\omega (=D/N)$ are proportions of defectives in the lot.
 $P_o(n_1) = (1-p)^{n_1}$; $P_o^*(n_1) = (1-\omega)^{n_1-1}$. As N decreases, P_o and P_o^* decrease; other quantities are not changes.
(iii) Note the values shown do not depend on p .

| p_1 | p_1' | p | p' | ω | n_1 | PROBABILITY OF CORRECT CLASSIFICATION OF NONDEFECTIVE | EXPECTED PERCENT REDUCTION IN TESTS |
|-------|--------|------|------|----------|-------|---|-------------------------------------|
| 0.98 | 0.05 | 0.95 | 0.05 | 0.05 | 6 | 0.9870 | 55.7 |
| | | | | | 8 | 0.9835 | 51.2 |
| | | | | | 10 | 0.9805 | 47.7 |
| | | | | | 12 | 0.9771 | 43.9 |
| | | | | 0.1 | 6 | 0.9785 | 34.8 |
| | | | | | 8 | 0.9732 | 29.5 |
| | | | | | 10 | 0.9690 | 24.4 |
| | | | | | 12 | 0.9656 | 20.9 |
| | | | | 0.2 | 6 | 0.9662 | 9.7 |
| | | | | | 8 | 0.9608 | 5.1 |
| | | | | | 10 | 0.9572 | 2.0 |
| | | | | | 12 | 0.9550 | 0.0 |

Table 1: Simple Screening (cont'd)

| P_1 | P_1' | p | p' | ω | n_1 | PROBABILITY OF CORRECT CLASSIFICATION OF NONDEFECTIVE | EXPECTED PERCENT REDUCTION IN TISTS |
|-------|--------|------|------|----------|-------|--|--|
| 0.95 | 0.05 | 0.95 | 0.05 | 0.05 | 6 | 0.9873 | 54.5 |
| | | | | | 8 | 0.9839 | 52.2 |
| | | | | | 10 | 0.9804 | 48.9 |
| | | | | | 12 | 0.9781 | 45.3 |
| | | | | 0.1 | 6 | 0.9791 | 56.2 |
| | | | | | 8 | 0.9740 | 51.2 |
| | | | | | 10 | 0.9699 | 46.4 |
| | | | | | 12 | 0.9666 | 42.1 |
| | | | | 0.2 | 6 | 0.9672 | 41.9 |
| | | | | | 8 | 0.9619 | 37.6 |
| | | | | | 10 | 0.9589 | 34.7 |
| | | | | | 12 | 0.9564 | 32.9 |

Table 1: Simple Screening (cont'd)

| P_1 | P_1' | P | P' | ω | n_1 | PROBABILITY OF CORRECT CLASSIFICATION OF NONDEFECTIVES | EXPECTED PERCENT REDUCTION IN TESTS ^o |
|-------|--------|------|------|----------|-------|---|---|
| 0.90 | 0.05 | 0.95 | 0.05 | 0.05 | 6 | 0.9878 | 55.8 |
| | | | | | 8 | 0.9847 | 53.9 |
| | | | | | 10 | 0.9818 | 50.9 |
| | | | | | 12 | 0.9792 | 47.6 |
| | | | | 0.1 | 6 | 0.9801 | 38.5 |
| | | | | | 8 | 0.9753 | 34.1 |
| | | | | | 10 | 0.9715 | 29.6 |
| | | | | | 12 | 0.9685 | 25.7 |
| | | | | 0.2 | 6 | 0.9689 | 15.5 |
| | | | | | 8 | 0.9639 | 11.8 |
| | | | | | 10 | 0.9607 | 9.1 |
| | | | | | 12 | 0.9587 | 7.5 |

Table 1: Simple Screening (cont'd)

| P_1 | P_1' | P | P' | w | n_1 | PROBABILITY OF CORRECT CLASSIFICATION OF NONDEFECTIVE | EXPECTED PERCENT REDUCTION IN TESTS |
|-------|--------|------|------|------|-------|--|--|
| 0.95 | 0.10 | 0.90 | 0.10 | 0.05 | 6 | 0.9718 | 60.8 |
| | | | | | 8 | 0.9644 | 58.9 |
| | | | | | 10 | 0.9586 | 55.9 |
| | | | | | 12 | 0.9533 | 52.6 |
| | | | | | 6 | 0.9552 | 43.5 |
| | | | | | 8 | 0.9457 | 39.1 |
| | | | | | 10 | 0.9370 | 34.6 |
| | | | | | 12 | 0.9317 | 30.7 |
| | | | | | 6 | 0.9329 | 20.5 |
| | | | | | 8 | 0.9228 | 16.8 |
| | | | | | 10 | 0.9164 | 14.1 |
| | | | | | 12 | 0.9123 | 12.5 |

Note that in the last set, the expected percent reduction in expected number of tests is always 5% greater than for the corresponding case in the penultimate set. This is because the value of $P_1 - P_1'$ is the same (0.85) in the two sets while the value of P_1 is 0.05 greater in the last set.

Table 2: Two-Stage Hierarchical Screening

| P_1 | P_1' | P_2 | P_2' | P' | ω | n_1 | n_2 | PROBABILITY OF CORRECT CLASSIFICATION OF NONDEFECTIVES | EXPECTED PERCENT REDUCTION IN TESTS |
|-------|--------|-------|--------|------|----------|-------|-------|---|--|
| 0.98 | 0.05 | 0.98 | 0.05 | 0.05 | 0.05 | 6 | 2 | 0.9971 | 65.2 |
| | | | | | | 6 | 5 | 0.9949 | 63.7 |
| | | | | | | 12 | 2 | 0.9966 | 56.5 |
| | | | | | | 12 | 5 | 0.9944 | 60.5 |
| | | | | | | 12 | 1 | 0.9921 | 60.4 |
| | | | | | | 12 | 6 | 0.9886 | 57.1 |
| | | | | | | | | | |
| | | | | | 0.1 | 6 | 2 | 0.9945 | 59.5 |
| | | | | | | 6 | 5 | 0.9902 | 40.6 |
| | | | | | | 12 | 2 | 0.9937 | 54.9 |
| | | | | | | 12 | 5 | 0.9896 | 59.5 |
| | | | | | | 12 | 1 | 0.9859 | 38.8 |
| | | | | | | 12 | 6 | 0.9796 | 35.4 |
| | | | | | | | | | |
| | | | | | 0.2 | 6 | 2 | 0.9892 | 10.1 |
| | | | | | | 6 | 5 | 0.9819 | 10.6 |
| | | | | | | 12 | 2 | 0.9886 | 8.5 |
| | | | | | | 12 | 5 | 0.9813 | 12.1 |
| | | | | | | 12 | 4 | 0.9755 | 10.4 |
| | | | | | | 12 | 6 | 0.9671 | 4.6 |

Table 2: Two-Stage Hierarchal Screening (cont'd)

| p_1 | p_1' | p_2 | p_2' | p' | ω | n_1 | n_2 | PROBABILITY OF CORRECT CLASSIFICATION OF NONDEFECTIVES | EXPECTED PERCENT REDUCTION IN TESTS(%) |
|-------|--------|-------|--------|------|----------|-------|-------|---|---|
| 0.98 | 0.05 | 0.90 | 0.05 | 0.05 | 0.05 | 6 | 2 | 0.9973 | 58.9 |
| | | | | | | 6 | 3 | 0.9959 | 60.1 |
| | | | | | | 12 | 2 | 0.9968 | 57.5 |
| | | | | | | 12 | 3 | 0.9948 | 61.5 |
| | | | | | 0.1 | 6 | 2 | 0.9948 | 40.8 |
| | | | | | | 6 | 3 | 0.9910 | 42.1 |
| | | | | | | 12 | 2 | 0.9911 | 36.4 |
| | | | | | | 12 | 3 | 0.9904 | 41.6 |
| | | | | | 0.2 | 6 | 2 | 0.9900 | 12.9 |
| | | | | | | 6 | 3 | 0.9855 | 14.5 |
| | | | | | | 12 | 2 | 0.9894 | 11.3 |
| | | | | | | 12 | 3 | 0.9828 | 15.9 |

Table 2: Two-Stage Hierarchical Screening (cont'd)

| P_1 | P_1' | P_2 | P_2' | P' | ω | n_1 | n_2 | PROBABILITY OF CORRECT CLASSIFICATION FOR NONDEFECTIVES | EXPECTED PERCENT REDUCTION IN TESTS |
|-------|--------|-------|--------|------|----------|-------|-------|--|--|
| 0.95 | 0.05 | 0.95 | 0.05 | 0.05 | 0.05 | 6 | 2 | 0.9972 | 59.2 |
| | | | | | | 6 | 3 | 0.9954 | 60.2 |
| | | | | | | 12 | 2 | 0.9968 | 57.8 |
| | | | | | | 12 | 3 | 0.9947 | 61.2 |
| | | | | | 0.1 | 6 | 2 | 0.9947 | 40.9 |
| | | | | | | 6 | 3 | 0.9908 | 42.0 |
| | | | | | | 12 | 2 | 0.9941 | 37.2 |
| | | | | | | 12 | 3 | 0.9902 | 41.8 |
| | | | | | 0.2 | 6 | 2 | 0.9898 | 15.4 |
| | | | | | | 6 | 3 | 0.9850 | 14.5 |
| | | | | | | 12 | 2 | 0.9895 | 11.9 |
| | | | | | | 12 | 3 | 0.9825 | 15.9 |

Table 2: Two-Stage Hierarchal Screening (cont'd)

| P_1 | P_1' | P_2 | P_2' | P' | ω | n_1 | n_2 | PROBABILITY OF CORRECT CLASSIFICATION OF NONDEFECTIVES | EXPECTED PERCENT REDUCTION IN TESTS |
|-------|--------|-------|--------|------|----------|-------|-------|---|--|
| 0.95 | 0.05 | 0.95 | 0.05 | 0.05 | 0.05 | 6 | 2 | 0.9972 | 59.2 |
| | | | | | | 6 | 5 | 0.9954 | 60.2 |
| | | | | | | 12 | 2 | 0.9968 | 57.8 |
| | | | | | | 12 | 5 | 0.9947 | 61.7 |
| | | | | | 0.1 | 6 | 2 | 0.9917 | 40.9 |
| | | | | | | 6 | 5 | 0.9908 | 42.0 |
| | | | | | | 12 | 2 | 0.9941 | 37.2 |
| | | | | | | 12 | 5 | 0.9902 | 41.8 |
| | | | | | 0.2 | 6 | 2 | 0.9898 | 15.4 |
| | | | | | | 6 | 5 | 0.9850 | 14.3 |
| | | | | | | 12 | 2 | 0.9895 | 11.9 |
| | | | | | | 12 | 5 | 0.9825 | 15.9 |

Table 2: Two-Stage Hierarchical Screening (cont'd)

| p_1 | p_1' | p_2 | p_2' | p | ω | n_1 | n_2 | PROBABILITY OF CORRECT CLASSIFICATION OF NONDEFECTIVES | EXPECTED PERCENT REDUCTION IN TESTS(%) |
|-------|--------|-------|--------|------|----------|-------|-------|---|---|
| 0.95 | 0.10 | 0.95 | 0.05 | 0.05 | 0.25 | 6 | 2 | 0.9971 | 57.2 |
| | | | | | | 6 | 3 | 0.9953 | 58.8 |
| | | | | | | 12 | 2 | 0.9967 | 56.3 |
| | | | | | | 12 | 3 | 0.9946 | 60.7 |
| | | | | | 0.1 | 6 | 2 | 0.9946 | 59.4 |
| | | | | | | 6 | 3 | 0.9907 | 41.0 |
| | | | | | | 12 | 2 | 0.9941 | 36.4 |
| | | | | | | 12 | 3 | 0.9902 | 41.3 |
| | | | | | 0.2 | 6 | 2 | 0.9898 | 12.7 |
| | | | | | | 6 | 3 | 0.9830 | 13.8 |
| | | | | | | 12 | 2 | 0.9893 | 11.7 |
| | | | | | | 12 | 3 | 0.9825 | 15.8 |

Note that in the last set the probabilities of correct classification of nondefectives are very slightly greater (no more than 0.0001) than the corresponding probabilities for the penultimate set, while the expected reduction in number of tests is slightly (c. 1-2%) less. (This comparison reflects the effect of changing p_1' from 0.05 to 0.10).

APPENDIX

In order to study the distribution of the total number (Z) of items declared "defective" we introduce the auxiliary indicator variables:

$$w_j = \begin{cases} 1 & \text{if items in the } j\text{-th subsample are tested individually} \\ 0 & \text{otherwise} \end{cases} \quad (A1)$$

for $j = 1, \dots, h$. Conditional on w_1, \dots, w_h and the *actual* numbers Y_1, \dots, Y_h of defectives in the corresponding subsamples, Z is distributed as

$$\text{Binomial} \left(\sum_{j=1}^h w_j Y_j, p \right) * \text{Binomial} \left(\sum_{j=1}^h w_j (n_2 - Y_j), p' \right) \quad (A2)$$

The conditional r -th factorial moment of Z is

$$E(Z^{(r)} | \underline{w}, \underline{Y}) = \sum_{i=0}^r \binom{r}{i} \left(\sum_{j=1}^h w_j Y_j \right)^{(i)} \left\{ \sum_{j=1}^h w_j (n_2 - Y_j) \right\}^{(r-i)} p^i p'^{r-i} \quad (A3)$$

Conditional on \underline{Y} for any $\alpha, \alpha' > 0$; and $j \neq j'$

$$E(w_j^\alpha | \underline{Y}) = P(w_j = 1 | \underline{Y}) = \begin{cases} p_1 p_2 & \text{if } Y_j = 0 \\ p_1' p_2' & \text{if } Y_j = 0, \sum_{i=1}^h Y_i > 0 \\ p_1' p_2' & \text{if } Y_j = 0, \sum_{i=1}^h Y_i > 0 \end{cases} \quad (A4)$$

$$E(w_j^\alpha w_{j'}^{\alpha'} | \underline{Y}) = P(w_j = 1, w_{j'} = 1 | \underline{Y}) = \begin{cases} p_1 p_2^2 & \text{if } Y_j = 0, Y_{j'} = 0 \\ p_1 p_2 p_2' & \text{if } Y_j > 0, Y_{j'} = 0 \text{ or } Y_j = 0, Y_{j'} > 0 \\ p_1' p_2'^2 & \text{if } Y_j = Y_{j'} = 0, \sum_{i=1}^h Y_i > 0 \\ p_1' p_2'^2 & \text{if } Y_j = Y_{j'} = 0 = \sum_{i=1}^h Y_i \end{cases} \quad (A5)$$

We have

$$P(Y_j = 0) = P_0(n_2); P(Y_j > 0) = 1 - P_0(n_2) \quad (A6.1)$$

$$P(Y_j = 0, Y_j' = 0) = P_0(2n_2) \quad (A6.2)$$

$$P(Y_j = 0, Y_j' > 0) = P(Y_j' > 0, Y_j = 0) = P_0(n_2) - P_0(2n_2) \quad (A6.3)$$

$$P(Y_j > 0, Y_j' > 0) = 1 - 2P_0(n_2) + P_0(2n_2) \quad (A6.4)$$

$$P(Y_j = Y_j' = 0, \sum_{i=1}^h Y_i > 0) = P_0(2n_2) - P_0(n_1) \quad (A6.5)$$

$$P(Y_j = 0 = \sum_{i=1}^h Y_i) = P(Y_j = Y_j' = 0 = \sum_{i=1}^h Y_i) = P_0(n_1) \quad (A6.6)$$

whence

$$E(W_j^\alpha) = p_1 p_2 - p_{(2)} p_1 - p_{(1)} p_2' \quad (A7.1)$$

$$E(W_j^\alpha W_j^{\alpha'}) = p_1 p_2^2 - 2p_1 p_2 p_{(2)} + p_1 (p_2 - p_2')^2 P_0(2n_2) - p_2'^2 p_{(1)} \quad (A7.2)$$

Also, from (A4) and (A5), with $\beta, \beta' > 0$

$$E(W_j^\alpha Y_j^\beta) = p_1 p_2 E(Y_j^\beta) \quad (A8.1)$$

$$E(W_j^\alpha W_j^{\alpha'} Y_j^\beta Y_j^{\beta'}) = p_1 p_2^2 E(Y_j^\beta Y_j^{\beta'}) \quad (A8.2)$$

$$E(W_j^\alpha W_j^{\alpha'} Y_j^\beta) = p_1 p_2^2 \{1 - P_0(n_2)\} - E(Y_j^\beta | Y_j' > 0) + p_1 p_2 p_2' P_0(n_2) E(Y_j^\beta | Y_j' = 0) \quad (A8.3)$$

The joint distribution function of \underline{Y} is

$$P(\underline{Y}) = \quad (A9)$$

$$\left[\begin{matrix} D \\ Y_1, \dots, Y_h, D - \sum_{i=1}^h Y_i \end{matrix} \right] \left[\begin{matrix} N-D \\ n_2 - Y_1, \dots, n_2 - Y_h, N-D-n_1 + \sum_{i=1}^h Y_i \end{matrix} \right] / \left[\begin{matrix} N \\ n_2, \dots, n_2, N-n_1 \end{matrix} \right]$$

and the h -th joint factorial moment of \underline{Y} is

$$E \left[\prod_{j=1}^h Y_j^{(r_j)} \right] = \left[\prod_{j=1}^h n_2^{(r_j)} \right] p^{(r)} / N^{(r)} \quad (A10)$$

where $r = \sum_{i=1}^h r_i$. In particular for all $j \neq j'$

$$E(Y_j) = n_2 D / N \quad (A11.1)$$

$$E(Y_j^2) = n_2 D (n_2 - 1) D + N - n_2^2 / N^{(2)} \quad (A11.2)$$

$$E(Y_j Y_{j'}) = n_2^2 D^{(2)} / N^{(2)} \quad (A11.3)$$

$$E(Y_j | Y_{j'} = 0) = n_2 D / (N - n_2); \quad 1(Y_j | Y_{j'} = 0) = n_2 D (1 - n_2 N^{-1} - p_0(n_2)) / (1 - p_0(n_2))^{-1} \times (N n_2)^{-1} \quad (A11.4)$$

Applying formulae (A5) - (A11) we get after some algebraic rearrangement the formula (22) for $E(Z)$. In order to calculate the variance of Z , we have to calculate

$$E(Z^{(2)}) = E \left[p^2 \left(\sum_{j=1}^h W_j Y_j \right) \left(\sum_{j=1}^h W_j Y_j - 1 \right) + 2 p p' \left(\sum_{j=1}^h W_j Y_j \right) \left(\sum_{j=1}^h W_j (n_2 - Y_j) \right) + p'^2 \left(\sum_{j=1}^h W_j (n_2 - Y_j) \right) \left(\sum_{j=1}^h W_j (n_2 - Y_j) - 1 \right) \right]$$

$$\begin{aligned}
 &= (p-p')^2 E\left[\left(\sum_{j=1}^h W_j Y_j\right)^2\right] - (p^2 - p'^2) E\left(\sum_{j=1}^h W_j Y_j\right) \\
 &\quad + 2n_2 p' (p-p') E\left[\left(\sum_{j=1}^h W_j Y_j\right) \left(\sum_{j=1}^h W_j\right)\right] - n_2 p'^2 E\left(\sum_{j=1}^h W_j\right) \\
 &\quad + n_2^2 p'^2 E\left[\left(\sum_{j=1}^h W_j\right)^2\right] \\
 &= (p-p')^2 [hE(W_j Y_j^2) + h(h-1)E(W_j W_j' Y_j Y_j')] - p^2 - p'^2 hE(W_j Y_j) \\
 &\quad + 2n_2 p' (p-p') [hE(W_j Y_j) + h(h-1)E(W_j W_j' Y_j)] - n_2 p'^2 hE(W_j) \\
 &\quad + n_2^2 p'^2 [hE(W_j) + h(h-1)E(W_j W_j')] \tag{A12}
 \end{aligned}$$

where $E(W_j)$ and $E(W_j W_j')$ are given by (A7), and

$$E(W_j Y_j) = n_2 p_1 p_2 D N^{-1} \tag{A15.1}$$

$$E(W_j Y_j^2) = n_2 p_1 p_2 D \{(n_2 - 1) D + N - n_2\} N^{-1} (N-1)^{-1} \tag{A15.2}$$

$$E(W_j W_j' Y_j) = n_2 p_1 p_2 D \{N^{-1} p_2 - (N - n_2)^{-1} p_{(2)}\} \tag{A15.3}$$

The resulting formulas for $E(Z^{(2)})$ and

$$\text{var}(Z) = E(Z^{(2)}) + E(Z) - \{E(Z)\}^2$$

are complicated, but numerical calculation is straightforward. If $p' = p_1' = p_2' = 0$ (so that there are no false positives), then

$$E(Z^{(2)}) = p^2 [hE(W_j Y_j^2) + h(h-1)E(W_j W_j' Y_j Y_j') - hE(W_j Y_j)] =$$

$$\frac{n_1 p_1 p_2 p^2 h}{N} \left[\frac{(n_2 - 1)(D - 1)}{N - 1} + (h - 1) \{p_2 - (1 - n_2 N^{-1}) p_{(2)}\} \right]$$

and also

$$E(Z) = \frac{n_1 p_1 p_2 p^D}{N} \quad (A14)$$

so that

$$\begin{aligned} \text{Var}(Z) = \frac{n_1 p_1 p_2 p^D}{N} & \left[\left\{ \frac{(n_2 - 1)(D - 1)}{N - 1} + (h - 1) [p_2 - (1 - n_2 N^{-1}) p_{(2)}] p \right. \right. \\ & \left. \left. + 1 - \frac{n_1 p_1 p_2 p^D}{N} \right\} \right] \quad (A15) \end{aligned}$$

NOTATION SUMMARY

Single Sampling

- p Probability that a defective item is classified as defective
 p' Probability that a nondefective item is classified as defective
 $\bar{p} = N^{-1}\{Dp + (N-D)p'\}$ = Probability that an individual chosen at random is classified as defective.

Hierarchical Sampling

- n_j Number of items tested *en bloc* at j -th stage
 p_j Probability that a group containing at least one defective is classified as defective, at j -th stage.
 p_j' Probability that a group containing no defectives is classified as defective, at j -th stage
 $P_0(n_j)$ = Probability that a random sample of n_j items contains no defective items ($j = 1, 2$).
 $P_0^*(n_j)$ = Probability that a random sample of n_j items contains no defective items, given that it contains at least one nondefective.
 $P_{(j)} = (p_j - p_j')P_0(n_j)$
 $P_{(j)}^* = (p_j - p_j')P_0^*(n_j)$

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|--|
| 1. REPORT NUMBER | 2. GOVT ACCESSION NO. AD-A111 219 | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) Errors in Inspection and Grading Distributional Aspects of Screening and Hierarchal Screening | | 5. TYPE OF REPORT & PERIOD COVERED TECHNICAL |
| 7. AUTHOR(s) Samuel Kotz and Norman L. Johnson | | 6. PERFORMING ORG. REPORT NUMBER |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS | | 8. CONTRACT OR GRANT NUMBER(s) ONR Contract N00014-81-K-0501 NSF Grant MCS-8-21704 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Office of Naval Research Statistics and Probability Program (Code 436) | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | | 12. REPORT DATE February 1982 |
| | | 13. NUMBER OF PAGES 27 |
| | | 15. SECURITY CLASS. (of this report) UNCLASSIFIED |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |
| 16. DISTRIBUTION STATEMENT (of this Report) Approved for Public Release -- Distribution Unlimited | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) | | |
| 18. SUPPLEMENTARY NOTES | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) hierarchal group screening; binomial distribution; compound distributions; faulty identification; hypergeometric distribution; sampling inspection; incomplete identification; stratified populations. | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) → Continuing the studies of Johnson <i>et al.</i> (1980) and Johnson and Kotz (1981), further distributions arising from models of errors in inspection and grading of samples from finite, possibly stratified lots are obtained. Screening, and hierarchal screening forms of inspection are also considered, and the effects of errors on the advantages of these techniques assessed. ← | | |

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

**DAI
ILMI**